sciendo

BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **20**, No 5
Special issue on Innovations in Intelligent Systems and Applications
Sofia • 2020 Print ISSN: 1311-9702; Online ISSN: 1314-4081
DOI: 10.2478/cait-2020-0040

Offline Signature Identification and Verification Based on Capsule Representations

Dilara Gumusbas, Tulay Yildirim

Electronics and Communication Engineering, Yildiz Technical University, Istanbul, Turkey
E-mails: f0415058@std.yildiz.edu.tr tulay@yildiz.edu.tr

Abstract: *Offline signature is one of the frequently used biometric traits in daily life and yet skilled forgeries are posing a great challenge for offline signature verification. To differentiate forgeries, a variety of research has been conducted on hand-crafted feature extraction methods until now. However, these methods have recently been set aside for automatic feature extraction methods such as Convolutional Neural Networks* (*CNN*). *Although these CNN-based algorithms often achieve satisfying results, they require either many samples in training or pre-trained network weights. Recently, Capsule Network has been proposed to model with fewer data by using the advantage of convolutional layers for automatic feature extraction. Moreover, feature representations are obtained as vectors instead of scalar activation values in CNN to keep orientation information. Since signature samples per user are limited and feature orientations in signature samples are highly informative, this paper first aims to evaluate the capability of Capsule Network for signature identification tasks on three benchmark databases. Capsule Network achieves 97 96, 94 89, 95 and 91% accuracy on CEDAR, GPDS-100 and MCYT databases for 64×64 and 32×32 resolutions, which are lower than usual, respectively. The second aim of the paper is to generalize the capability of Capsule Network concerning the verification task. Capsule Network achieves average 91, 86, and 89% accuracy on CEDAR, GPDS-100 and MCYT databases for 64×64 resolutions, respectively. Through this evaluation, the capability of Capsule Network is shown for offline verification and identification tasks.*

Keywords: *Capsule Network, Offline Signature Verification, Offline Signature Identification, Convolutional Neural Networks.*

1. Introduction

Biometrics is a field that uses behavioral and biological traits to identify/verify a person. Some of these traits are fingerprint, iris, gait and signature. Due to ease of collecting and being non-invasive, signature-based biometric systems are frequently

used. These systems are divided into two depending on their collection method: online and offline. The first collects dynamic signature features as a sequence of time. In this manner, features such as speed, pressure can be extracted. The second uses the image after signing is done. Even though online signature is advantageous for keeping more details about a signature, the offline signature is the most frequently used behavioral trait in daily life [1].

In offline signature verification systems, the main goal is to differentiate the genuine signatures from forgeries, which can be random, simple or skilled done by a forger. Unlike random and simple forgeries, skilled forgeries are not always easy to distinguish due to the intra-class variance of genuine signatures shown in Fig. 1. Therefore, a detailed investigation of not only local but also global features of genuine signatures is required to achieve high verification results. Moreover, insufficient prior knowledge about forgeries during training and limited genuine samples make the verification process even more challenging.

Fig. 1. Two genuine (first two rows) and one forgery signature (last row) samples from CEDAR, GPDS and MCYT databases, respectively [2-4]

Many research has been devoted to extracting the most informative global and, in particular local feature representations to differentiate forgeries. These hand-crafted local descriptors can be texture-based such as gray level co-occurrence matrix [5], direction-based such as Histogram of Gradients (HoG) [6], Scale Invariant Feature Transform (SIFT) [7] or combination of two or more different local descriptors [8]. While research on hand-crafted local descriptors is still in progress, recent researches have been conducted by employing automatic feature extraction algorithms such as Convolutional Neural Network (CNN). Since data samples per user are limited, a few studies are dedicated to using transfer learning instead of conducting data augmentation. The works represented in [9, 10] use a pre-trained CNN-based model after training the model with another benchmark dataset to make weight parameters hard-tuned. After hard-tuning, limited training data from the original dataset is used for fine-tuning. In brief, hard tuning is employed to narrow down signature feature space while fine tuning is to guarantee optimal decision boundaries. Similarly, the model proposed in [11] employs Siamese CNN with an inception layer. To cope with the few data samples per user, the model generates augmented samples for training. The model achieves 99.15 and 99.82 AUC rates for sub CEDAR and MCYT databases.

Unlike pre-trained CNN-based models, the method proposed in [12] narrows feature space down by only modeling with few data from the original dataset. The method, known as Capsule Network, asserts to obtain better feature representations than CNN due to its vector feature representations and inner affine transformation matrix. The main goal of this paper is to evaluate Capsule Network under different input resolutions, such as 64×64, 32×32, which are four to eight times lower than the usual signature resolutions for signature verification and identification tasks. This evaluation is chosen not only to investigate the detection capability of Capsule Network without requiring pre-trained weights under extremely low resolutions but also to fasten evaluation times and lessen memory usage.

The remainder of this paper is organized as follow:

- Section 2 outlines benchmark datasets and their preprocessing steps.
- Section 3 explains Capsule Network, evaluation procedures for identification and verification tasks and concludes with performance results.
- Section 4 discusses the results and points out future research directions.

2. Benchmark datasets and preprocessing steps

2.1. Benchmark datasets

In this paper, three frequently used offline signature datasets are employed for identification and verification tasks. This subsection includes the contents of these datasets in detail.

2.1.1. CEDAR dataset

CEDAR database consists of 1320 genuine and 1320 forgery samples in total and 24 genuine and 24 forgery samples are collected per user among 55 users [2].

2.1.2. MCYT dataset

MCYT database consists of 1125 genuine and 1125 forgery samples in total and 15 genuine and 15 simulated forgery samples are collected per user among 75 users [4].

2.1.3. GPDS dataset

GPDS database consists of 96,000 genuine and 120,000 forgery samples in total and 24 genuine and 30 simulated forgery samples are collected per user among 4000 users [3]. In this paper, we employed only the first 100 users for identification and verification tasks.

2.2. Preprocessing steps

Before the evaluation procedure, benchmark databases are preprocessed as shown in Fig. 2. First, data samples for each dataset are cropped regarding the center of signatures to discard unnecessary parts. Then, these data samples are resized to 64×64 and 32×32 extreme image resolutions. After resizing is done, data samples are

converted into binary values with Otsu's method. As a final step of preprocessing, binarizing is done to make background pixels black, foreground pixels white.

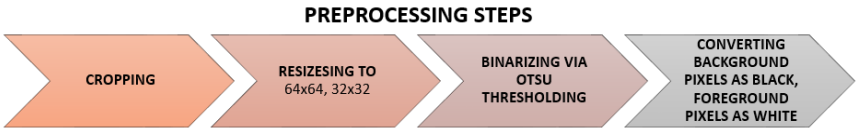

Fig. 2. Preprocessing steps for benchmark databases

3. Capsule Network and experimental setups

3.1. Capsule Network

Capsule Network is proposed to obtain more accurate representations of an object without requiring samples from different viewpoints [12]. The model structure is shown in Fig. 3. As can be seen from the figure, the model consists of four main parts, which are convolution layers, Primary Capsules, Signature Capsules and fully connected layers, respectively. Firstly, a variety of convolutions is applied to input images to extract local features in convolution layers. After getting activations as the output of these convolution layers, all these scalar-valued activations are given to primary capsules to be grouped into multi-dimensional vector representations.

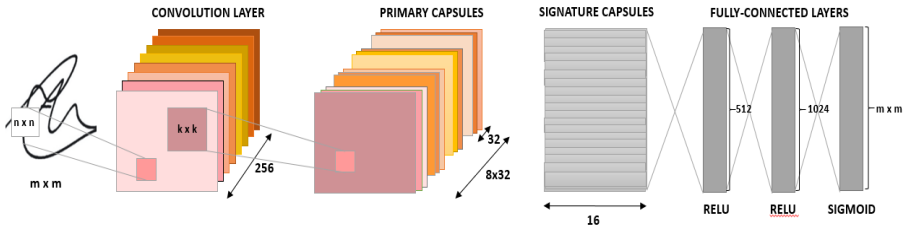

Fig. 3. Capsule Network model structure for Signature Identification and Verification systems

Then, these multi-dimensional vectors are multiplied with an affine transformation matrix to obtain many different variations of these vectors for better modeling. To select the most informative feature vectors, a routing algorithm is employed. Before employing this routing algorithm, all transformed feature vectors are squashed according to the next equation

$$v_j = \frac{\|s_j\|^2}{1+\|s_j\|^2} * \frac{s_j}{\|s_j\|^2} \quad (1)$$

to make discriminative feature vectors more apparent and to fade less-informative ones out. Here v_j is the output of a capsule, s_j is the total input of a capsule and it also includes affine transformed versions of convolution outputs which can be studied in detail from [12].

After squashing, the most informative vectors are routed to signature capsules to form an entity. When the routing algorithm is agreed, signature capsules with one

multidimensional capsule per class are created. The new capsule keeps the information about all outputs of capsules from the previous layer and keeps absolute characteristic features for each class.

Lastly, these signature capsules are fed into fully connected layers to classify. The loss function is designed as a combination of margin loss, which is obtained from false predictions, and reconstruction loss. This loss function is calculated according to

$$L = T_c \max(0, m^+ - \|v_c\|)^2 + \lambda(1 - T_c)\max(0, \|v_c\| - m^-)^2, \quad (2)$$

where L is the loss term for one signature capsule, T_c is a constant that is 1 if the signature capsule is the correct, else 0. First-term of loss equation is to calculate correct prediction probability while the latter term is employed to calculate incorrect prediction probability.

Overall, Capsule Network brings three main novelty compared to Convolutional Neural Network:

- Inner affine matrix multiplication instead of data augmentation,
- Vector representation instead of scalar-valued representation,
- Forwarding only the most informative feature representations by Dynamic Routing algorithm instead of forwarding all extracted feature activation values.

3.2. Experimental setups

General settings for the identification task are given below.

- One model is trained for all users in a specific dataset.
- Only genuine samples are used for training and testing. Train and test partitions are set as the first half and the second half of genuine samples per user, respectively.
- Two-fold cross-validation is employed.
- For training, epoch size and batch size are chosen as 50 and 16, respectively.
- Image resolutions of 64×64 and 32×32, which are 4-8 times lower than the usual, are used for identification tasks [9, 13-15].
- Model hyper-parameters such as routing number are chosen as the original in [12]. Only convolution kernel sizes and capsule dimensions are modified as given in Table 1.

General settings for the verification task are given below.

- A model is trained for each user separately, which is also known as the writer-dependent approach. For example, 55 separate models are created for 55 users in the CEDAR dataset and only average accuracy of all models is reported.
- Genuine as well as random and simple forgery samples, which are treated as a separate class, are used for training and testing. Train partition is set as the first half of genuine and forgery samples per user while the test partition is set as the remaining half.
- Two-fold cross-validation is employed.
- For training, epoch size and batch size are chosen as 50 and 16, respectively.

● Only image resolution 64×64 is used since the inner variance of genuine signatures makes modeling genuine signatures difficult for resolutions smaller such as 32×32.

● Model hyper-parameters such as routing number are chosen as the original in [12]. Capsule dimensions are set the same with identification tasks. Only convolution layers and kernel sizes are modified separately as given in Table 2 for each dataset.

Table 1. Test accuracies for offline signature identification task

Resolution	Dataset	$n \times n$(stride)	$k \times k$(stride)	Train	Test	Accuracy, %
64×64	CEDAR	21×21(1)	21×21(2)	12	12	97
32×32	CEDAR	13×13(1)	11×11(2)	12	12	96
64×64	GPDS-100	21×21(1)	21×21(2)	12	12	94
32×32	GPDS-100	13×13(1)	11×11(2)	12	12	89
64×64	MCYT	21×21(1)	21×21(2)	8	7	95
32×32	MCYT	13×13(1)	11×11(2)	8	7	91

Table 2. Test accuracies for offline signature verification task

Resolution	Dataset	$n \times n$(stride)	$k \times k$(stride)	Train	Test	Average Accuracy, %
64×64	CEDAR	21×21(1)	21×21(2)	14+14	5+5	91
64×64	GPDS-100	3×3(1)	5×5(2)	12+15	12+15	86
64×64	MCYT	21×21(1)	21×21(2)	8+8	7+7	89

3.3. Experimental results

This paper is an extended version of [16]. In addition to experiments done on CEDAR only, the paper increases the scope over several frequently used signature datasets such as MCYT and GPSD-100, which are in different complexity considering genuine and forgery samples per user. Therefore, the main goal of this paper is to generalize the capability of Capsule Network for signature identification and verification tasks.

All experimental results for offline signature identification tasks are given with the information of input resolutions, train-test partitions and convolution kernel sizes with stride in Table 1. As can be shown in Table 1, even big convolution kernels are good at modeling and separating signatures from one another. Additionally, identification at input resolutions of 32×32 achieves average 92% accuracy over three benchmark datasets.

All experimental results for verification tasks are given in Table 2. As can be seen in Table 2, only 64×64 image resolutions are used due to the difficulty level at differentiating forgeries from genuine.

Moreover, genuine and forgery samples in the GPDS-100 dataset seem almost identical. Therefore, they require smaller kernels to extract local features in detail for verification tasks. Only for this dataset, additional two convolution layers before primary capsules are added as 3×3(1) and 5×5(2). Additionally, train and test samples are given as genuine+forgery format.

4. Conclusion

Capsule Network acquires promising results while using at least four times lower resolution than frequently used ones for identification tasks. This indicates that Capsule Network is reliable enough to classify signatures and to have a unique ability to model local features better under extremely low resolutions. Moreover, results for identification tasks reveal that even using bigger sized (around one-third of input resolutions) convolutions are useful to separate signatures owing to the modeling capability of Capsule Network.

Similarly, verification results also indicate that the algorithm has a great capability to cope with differentiating genuine signatures from forgeries. However, unlike identification tasks, high-similarity between genuine and forgery samples requires the extraction of low and mid-level features together. Moreover, different benchmark datasets make usage of different levels of convolution layers necessary. For instance, two-layer kernels are enough to extract enough information in CEDAR and MCYT datasets while GPDS-100 dataset requires more. Therefore, convolution layers and convolution kernel sizes are arranged for that requirement for all datasets.

For future works, there are a couple of things to be taken into consideration for offline signature identification and verification tasks. These are:

- Different model combinations for Capsule Network may be tried, such as modifying the stride, convolution layers, etc. Lastly, data augmentation can be attempted.
- Capsule Network and state-of-art CNN models such as VGG-16 can be compared for high input resolutions using the same convolution layers to investigate modeling capability of Capsule Network in detail.
- For verification tasks, performance comparison can be generalized with the use of adversarial attacks, such as adding noise to genuine signatures.
- Visualizing feature representations before and after capsule layers can be done to increase the explainability of capsule-based feature modeling.
- Evaluation metrics such as Equal Error Rate (EER), DET curves can be employed.

In conclusion, the main point of this paper is to show Capsule Network's advantages in terms of data representation, using less data for signature identification and verification tasks and to encourage a community that is interested in online signature verification to think one step further to obtain better feature representations for the future.

References

1. Hafemann, L. G., R. Sabourin, L. S. Oliveira. Offline Handwritten Signature Verification: Literature Review. – In: Proc. of 7th International Conference on Image Processing Theory, Tools and Applications (IPTA'17), November 2017, pp. 1-8.
2. Kalera, M. K., S. N. Srihari, A. Xu. Offline Signature Verification and Identification Using Distance Statistics. – IJPRAI, Vol. **18**, 2004, pp. 1339-1360.
3. Vargas, F., M. Ferrer, C. Travieso, J. Alonso. Off-Line Handwritten Signature Gpds-960 Corpus. –In: Proc. of 9th International Conference on Document Analysis and Recognition (ICDAR'07), Vol. **2**, September 2007, pp. 764-768.
4. Ortega-Garcia, J. Fierrez-Aguilar, D. Simon, J. Gonzalez, M. Faundez-Zanuy, V. Espinosa, A. Satue, I. Hernaez, J. Igarza, C. Vivaracho, D. Escudero, Q. Moro. MCYT Baseline Corpus: A Bimodal Biometric Database. – IEE Proceedings – Vision, Image and Signal Processing, Vol. **150**, December 2003, No 6, pp. 395-401.
5. Hiryanto, A., R. Yohannis, T. Handhayani. Hand Signature and Handwriting Recognition as Identification of the Writer Using Gray Level Cooccurrence Matrix and Bootstrap. – In: Proc. of 2017 Intelligent Systems Conference (IntelliSys'17), September 2017, pp. 1103-1110.
6. Yilmaz, M. B., B. Yanikoglu. Score Level Fusion of Classifiers in Off-Line Signature Verification. – Information Fusion, Vol. **32**, 2016.
7. Ruiz-Del-Solar, J., C. Devia, P. Loncomilla, F. Concha. Offline Signature Verification Using Local Interest Points and Descriptors. – In: Proc. of 13th Iberoamerican Congress on Pattern Recognition: Progress in Pattern Recognition, Image Analysis and Applications, Ser. CIARP'08, Berlin, Heidelberg, Springer-Verlag, 2008, p. 22-29.
8. Yilmaz, M. B., B. Yanikoglu, C. Tirkaz, A. Kholmatov. Offline Signature Verification Using Classifier Combination of Hog and Lbp Features. – In: Proc. of 2011 International Joint Conference on Biometrics (IJCB'11), October 2011, pp. 1-7.
9. Hafemann, L. G., R. Sabourin, L. S. Oliveira. Learning Features for Offline Handwritten Signature Verification Using Deep Convolutional Neural Networks. – Pattern Recognition, Vol. **70**, 2017, pp. 163-176.
10. Yilmaz, M. B., K. Ozturk. Hybrid User-Independent and User-Dependent Offline Signature Verification with a Two-Channel CNN. – In: Proc. of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'18), 2018, pp. 639-6398.
11. Ruiz, V., I. Linares, A. Sanchez, J. F. Velez. Off-Line Handwritten Signature Verification Using Compositional Synthetic Generation of Signatures and Siamese Neural Networks. – Neurocomputing, Vol. **374**, 2020, pp. 30-41.
12. Sabour, S., N. Frosst, G. E. Hinton. Dynamic Routing Between Capsules. – CoRR, Vol. **abs/1710.09829**, 2017.
    http://arxiv.org/abs/1710.09829
13. Hafemann, L. G., R. Sabourin, L. Oliveira. Characterizing and Evaluating Adversarial Examples for Offline Handwritten Signature Verification. – IEEE Transactions on Information Forensics and Security, 2019, pp. 1-1.
14. Mersa, O., F. Etaati, S. Masoudnia, B. N. Araabi. Learning Representations from Persian Handwriting for Offline Signature Verification, a Deep Transfer Learning Approach. arXiv e-prints, February 2019.
15. Souza, V. L. F., A. L. I. Oliveira, R. Sabourin. A Writer-Independent Approach for Offline Signature Verification Using Deep Convolutional Neural Networks Features. – In: Proc. of 7th Brazilian Conference on Intelligent Systems (BRACIS'18), October 2018, pp. 212-217.
16. Gumusbas, D., T. Yildirim. Offline Signature Identification and Verification Using Capsule Network. – In: Proc. of IEEE International Symposium on INnovations in Intelligent SysTems and Applications (INISTA'19), July 2019, pp. 1-5.

Received: 21.02.2020; Second Version: 05.05.2020; Accepted: 08.05.2020